

# SCIENTIFIC REPORTS

OPEN

## *Xylella fastidiosa*: climate suitability of European continent

Martin Godefroid, Astrid Cruaud, Jean-Claude Streito, Jean-Yves Rasplus & Jean-Pierre Rossi

The bacterium *Xylella fastidiosa* (*Xf*) is a plant endophyte native to the Americas that causes diseases in many crops of economic importance (grapevine, *Citrus*, Olive trees etc). *Xf* has been recently detected in several regions outside of its native range including Europe where little is known about its potential geographical expansion. We collected data documenting the native and invaded ranges of the *Xf* subspecies *fastidiosa*, *pauca* and *multiplex* and fitted bioclimatic species distribution models (SDMs) to assess the potential climate suitability of European continent for those pathogens. According to model predictions, the currently reported distribution of *Xf* in Europe is small compared to the large extent of climatically suitable areas. The regions at high risk encompass the Mediterranean coastal areas of Spain, Greece, Italy and France, the Atlantic coastal areas of France, Portugal and Spain as well as the southwestern regions of Spain and lowlands in southern Italy. The extent of predicted climatically suitable conditions for the different subspecies are contrasted. The subspecies *multiplex*, and to a certain extent the subspecies *fastidiosa*, represent a threat to most of Europe while the climatically suitable areas for the subspecies *pauca* are mostly limited to the Mediterranean basin. These results provide crucial information for the design of a spatially informed European-scale integrated management strategy, including early detection surveys in plants and insect vectors and quarantine measures.

The bacterium *Xylella fastidiosa* (*Xf*) is a plant endophyte native to the Americas, which develops in up to 300 plant species including ornamental and agricultural plants<sup>1</sup>. *Xf* is transmitted between plants by xylem-feeding insects belonging to several families of Hemiptera (Aphrophoridae, Cercopidae, Cicadellidae, Cicadidae and Clastopteridae)<sup>2</sup>. *Xf* causes severe plant pathologies leading to huge economic losses<sup>3</sup> e.g., the Pierce's disease of grapevines PD<sup>4</sup>, the olive quick decline<sup>5</sup>, the oak bacterial leaf scorch<sup>6</sup>, the phony peach disease<sup>7</sup>, the *Citrus* variegated chlorosis CVC<sup>8</sup> and the almond leaf scorch<sup>9</sup>. As *Xf* induces diseases to a large number of economically important plants including vine<sup>1</sup>, the biology of this pathogen and the mechanisms of vector transmission have been extensively studied to design management strategies<sup>10</sup>.

On the basis of genetic data obtained with Multilocus Sequence Typing MLST<sup>11,12</sup>, *Xf* was subdivided into six subspecies (*fastidiosa*, *morus*, *multiplex*, *pauca*, *sandyi* and *tashke*). Those subspecies were further characterized by different geographic origins, distributions and host preferences in the Americas<sup>13–15</sup>. However, the intraspecific taxonomic boundaries of *Xf* are still debated<sup>16</sup> and only the two subspecies *fastidiosa* and *multiplex* are formally considered valid names<sup>1,17</sup>. *Xf* subsp. *fastidiosa*<sup>18</sup> occurs in North and Central America, where it causes, among others, the harmful PD and the almond leaf scorch (ALS). Genetic analyses suggest that this subspecies originates from southern parts of Central America<sup>19</sup>. The subspecies *multiplex* is widely distributed in North America (from California to western Canada and from Florida to eastern Canada), where it was detected on a wide range of host plants (e.g., oak, elm, maple, almond, sycamore, *Prunus* sp., etc.) as well as in South America<sup>20,21</sup>. The subspecies *pauca*, which causes severe diseases in *Citrus* (CVC) and coffee (Coffee Leaf Scorch)<sup>22</sup> in South and Central America, is speculated to be native to South America<sup>23</sup>. The subspecies *morus* recently proposed by Nunney *et al.*<sup>24</sup>, occurs in California and eastern USA, where it is associated to mulberry leaf scorch. *Xf* subsp. *sandyi*, responsible for oleander leaf scorch, is distributed in California<sup>12</sup>, while the subspecies *tashke* was proposed by Randall *et al.*<sup>25</sup> for a strain occurring on *Chitalpa tashkentensis* in New Mexico and Arizona. Overall, intraspecific entities of *Xf* display noticeable differences in host range suggesting that the radiation of *Xf* into multiple subspecies and strains is primarily associated to host specialization<sup>26</sup>.

*Xylella fastidiosa* is now of worldwide concern. In 2013, the CoDIRO strain (subsp. *pauca*) was detected on olive trees in southern part of the Apulia territory (Italy). Genetic analyses suggest that this strain was accidentally introduced in Italy from Costa Rica or Honduras via infected ornamental coffee plants<sup>5</sup>. Since then, *Xf* subsp.

CBGP, INRA, CIRAD, IRD, Montpellier SupAgro, Montpellier, France. Correspondence and requests for materials should be addressed to J.-P.R. (email: [jean-pierre.rossi@inra.fr](mailto:jean-pierre.rossi@inra.fr))

*pauca* has spread northward and killed millions of olive trees in the Apulia territory, causing unprecedented socio-economic issues. During the period 2015–2017 several subspecies and strains were detected on ca. 30 different host plants in Southern France (PACA region) and Corsica<sup>27</sup>. According to national surveys performed in France, the vast majority of plant samples were contaminated by two strains of *Xf* subsp. *multiplex*, and two strains were identified (hereafter referred to as the French ST6 and ST7 strains)<sup>27</sup>. These strains are closely related to the Californian strains Dixon (ST6) and Griffin (ST7), belonging to the “almond group”<sup>26</sup> and that were detected on numerous plant species, though without evident specialization. To a lesser extent, other strains occur in Southern France i.e., the strain ST53 (*Xf* subsp. *pauca*) was detected on *Polygala myrtifolia* in Côte d’Azur (Menton) and on *Quercus ilex* in Corsica<sup>27</sup> and the recombinant strains (ST76, ST79 or not yet fully characterized) were detected in a few plant samples. In 2016, *Xf* subsp. *fastidiosa* was detected on rosemary and oleander plants overwintering in a nursery in Germany<sup>28</sup>. In 2017, Spanish plant biosecurity agencies officially confirmed the detection of *Xf* strains belonging to the subspecies *multiplex*, *pauca* and *fastidiosa* on almond trees, grapevine, cherry and plums in western parts of the Iberian Peninsula and Balearic islands<sup>29</sup>. Outside Europe, the detection of *Xf* was officially confirmed in Iran on almond trees and grapevines<sup>30</sup>, in Turkey<sup>31</sup> as well as in Taiwan on grapevines<sup>32</sup>.

The severity of *Xf*-induced diseases has recently increased possibly due to global warming<sup>33</sup>. Indeed, it has been demonstrated that cold winter temperatures might affect the survival of *Xf* in xylem vessels and allow plants to partly recover from *Xf*-induced diseases (‘cold curing phenomenon’)<sup>34,35</sup>. For instance, Purcell<sup>34</sup> showed that grapevines with symptoms of PD recovered after multiple exposures to temperatures below  $-8^{\circ}\text{C}$  during several hours. Further, Anas *et al.*<sup>36</sup> suggested that areas experiencing more than 2 to 3 days with minimal temperature below  $-12.2^{\circ}\text{C}$  (or alternatively 4 to 5 days below  $-9.9^{\circ}\text{C}$ ) should be considered at low risk for PD incidence, although these thresholds were considered too conservative by Lieth *et al.*<sup>37</sup>. Several studies aimed to forecast the potential distribution of *Xf* in Europe<sup>38</sup> and/or all over the world<sup>39</sup>. For instance, Hoddle *et al.*<sup>39</sup> used the CLIMEX algorithm to forecast the worldwide potential severity of PD. Their model suggested that most of Mediterranean areas are suitable for PD even though cold in winter would presumably hamper *Xf* range expansion into several of the most economically-important wine-producing regions of France and central and northern parts of Spain and Italy. Bosso *et al.*<sup>38</sup> fitted a Maxent model to forecast the potential distribution of *Xf* subsp. *pauca* under current and future climate conditions, and concluded that climate change would not affect the future distribution of *Xf*. Here, we analyze the potential distribution of three subspecies of *Xylella fastidiosa* using datasets describing native and newly established ranges and a set of four different species distribution models.

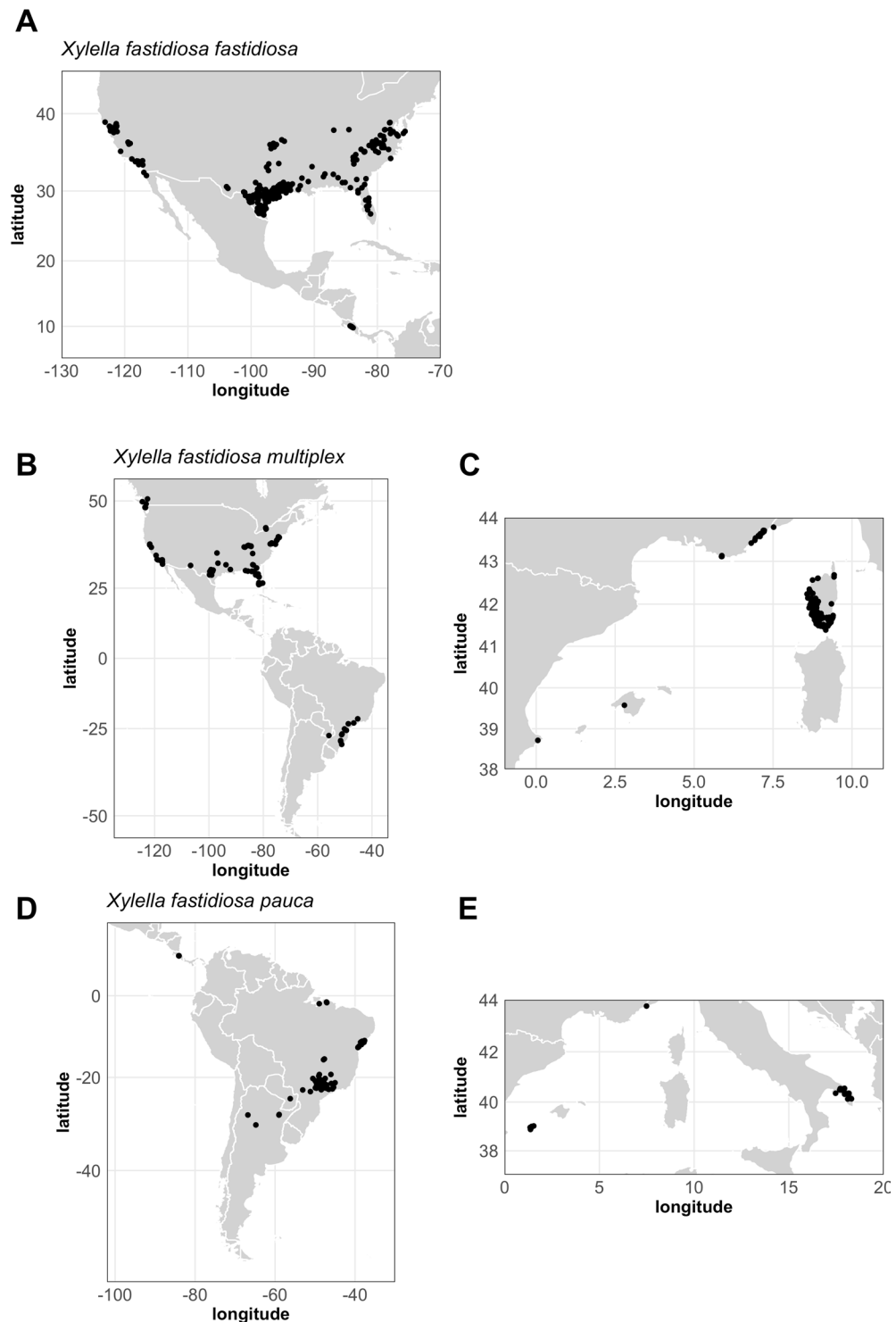
## Material and Methods

**Distribution data.** We collected occurrence data for subspecies *fastidiosa*, *multiplex* and *pauca* from the scientific literature, field surveys and public databases (Fig. 1). These datasets comprised both occurrences from native area (the Americas) and recently invaded regions (south Italy, France and Spain - Fig. 1). The occurrences located in France were collected in 2015–2019 and stored in the French national database managed by the French Agency for Food, Environmental and Occupational Health & Safety (ANSES) (Fig. 1C). For each subspecies, we randomly generated 10,000 background points within a wide area to properly depict the background environment in the Maxent calibration (see below)<sup>40,41</sup> (Supplementary Fig. S1). Pseudo-absences were randomly generated within regions located in the native area where we could confidently consider that the studied subspecies is absent. We restricted these areas to cold regions because low winter temperatures constitute a well-known factor constraining the distribution of *Xf*<sup>35</sup> (Supplementary Fig. S1).

**Bioclimatic descriptors.** We used a set of bioclimatic descriptors hosted in the Worldclim database<sup>42</sup>. We retained raster layers of 2.5-minute spatial resolution, which corresponds to about 4.5 km at the equator. The data represent the average climate conditions for the period 1970–2000.

**Models.** The potential distribution of *Xf* subsp. *fastidiosa*, *pauca* and *multiplex* were assessed using species distribution modeling. Such approach establishes mathematical species-environment relationships using occurrence/absence records and environmental descriptors in order to assess the potential distribution of species<sup>43</sup>. Different modeling techniques exist and their ability to predict species distribution / habitat suitability is known to vary substantially according to the algorithm used, the occurrence dataset, the environmental descriptors and the model calibration parameters<sup>44,45</sup>. As a consequence, there is no single ‘best’ modeling technique<sup>46</sup>. Such variation has led to the development of ensemble forecasting approach, which aims at building more robust forecasts by combining individual models in a consensus model<sup>47</sup>. However, averaging models’ outputs in the form of probabilities might raise issues. Indeed, models’ response to species prevalence could differ and yield non-comparable probabilities. A solution consists to use a committee averaging approach<sup>44,48</sup>.

In the present study, we adopted a two-step modeling strategy<sup>46</sup>. In the first step of the analysis we tested a set of four algorithms known for their good performance in species distribution modeling (Table 1). The test consisted in (i) fitting models using the occurrences available in the native areas only and (ii) evaluating the predictive power of those models into the recently colonized areas in Europe. It is known that bioclimatic descriptors used to calibrate the models can strongly impact performance and transferability. However, choosing proper descriptors is not easy. Consequently, we constituted seven subsets of bioclimatic descriptors by associating different variables that we *a priori* considered as ecologically meaningful when working on *Xf* (Table 1). We intentionally used a limited number of climate descriptors comprised between two to four to avoid model over-parameterization, which is a recommended practice, particularly when assessing invasion risk<sup>49</sup>. We used two descriptors of the low temperatures during the coldest periods of the year (bio6: minimum temperature of the coldest month; bio11: mean temperature of the coldest quarter). We also used a variable reflecting high temperatures during the warmest period of the year (bio10) and a variable reflecting the rainfall seasonality (bio15) that recently proved to be a good predictor of the spatial distribution of *Xf* in Corsica<sup>50</sup>. For each subspecies of *Xf*, algorithms were



**Figure 1.** Occurrences of three *Xylella fastidiosa* subspecies used in the study. (A) *Xylella fastidiosa fastidiosa*, (B) *Xylella fastidiosa multiplex* in its native range and (C) in Europe, (D) *Xylella fastidiosa pauca* in its native range and (E) in Europe. European occurrences of *Xylella fastidiosa fastidiosa* are sparse and not shown.

calibrated using each climate data subsets and the occurrences available in the Americas. The predictive power of each model was evaluated using the area under the curve of the receiver operating curve (AUC) statistic and the true skill statistic (TSS)<sup>51,52</sup>. This procedure aimed to identify the combinations of algorithms × climate datasets with weak predictive performances and to discard them from further analysis<sup>46</sup>.

In a second step of the analysis we used the best combinations of algorithms × climate datasets to calibrate models with all available subspecies occurrences to predict habitat suitability in Europe (referred to as full models). For each combination of model and climate datasets we calibrated the models using a randomly

Algorithm	metric	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7
		bio6	bio6	bio6	bio6	bio6		
		bio10	bio10	bio10	bio10		bio10	bio10
			bio11	bio11			bio11	bio11
				bio15	bio15	bio15		bio15
<b>Xf fastidiosa</b>								
Ann	AUC	0.93–0.96	0.94–0.99	0.98–1	0.98–0.99	0.98–0.98	0.89–0.94	0.98–0.99
—	TSS	0.78–0.85	0.81–0.92	0.91–0.95	0.86–0.93	0.78–0.9	0.63–0.82	0.86–0.91
bioclim	AUC	0.92–0.94	0.92–0.94	0.93–0.94	0.93–0.94	0.89–0.93	0.89–0.91	0.89–0.92
—	TSS	0.77–0.86	0.78–0.86	0.79–0.88	0.78–0.88	0.75–0.85	0.71–0.81	0.73–0.82
GLM	AUC	0.92–0.96	0.97–0.97	0.96–0.98	1–1	0.97–0.98	0.93–0.93	0.99–1
—	TSS	0.8–0.9	0.8–0.82	0.75–0.92	0.93–0.97	0.72–0.86	0.61–0.73	0.87–0.91
maxent	AUC	0.94–0.95	0.94–0.94	0.98–0.98	0.98–0.99	0.97–0.98	0.92–0.93	0.94–0.96
—	TSS	0.67–0.82	0.6–0.68	0.79–0.86	0.76–0.9	0.84–0.89	0.66–0.79	0.62–0.69
<b>Xf multiplex</b>								
Ann	AUC	0.99–1	0.99–1	1–1	1–1	1–1	1–1	0.99–1
—	TSS	0.94–0.97	0.96–0.99	0.95–1	0.99–1	1–1	1–1	0.94–1
bioclim	AUC	0.94–1	0.94–1	0.9–0.99	0.9–0.99	0.94–0.99	0.93–0.99	0.88–0.98
—	TSS	0.82–0.95	0.79–0.93	0.7–0.9	0.73–0.92	0.79–0.94	0.76–0.89	0.65–0.84
GLM	AUC	1–1	0.97–0.98	0.93–0.98	0.97–0.98	1–1	1–1	1–1
—	TSS	0.89–0.97	0.92–0.93	0.86–0.93	0.93–0.97	0.98–1	0.94–1	0.97–1
maxent	AUC	0.99–1	0.88–0.92	0.9–0.93	0.86–0.89	0.99–1	0.98–0.99	0.9–0.94
—	TSS	0.73–0.89	0.64–0.64	0.6–0.74	0.62–0.74	0.77–0.89	0.79–0.86	0.7–0.87
<b>Xf pauca</b>								
Ann	AUC	1–1	1–1	0.97–0.97	1–1	1–1	NC	0.97–1
—	TSS	0.99–1	0.99–1	0.94–0.94	1–1	0.99–1	NC	0.94–0.94
bioclim	AUC	0.92–1	NC	NC	0.92–1	1–1	NC	NC
—	TSS	0.85–1	NC	NC	0.85–1	0.98–1	NC	NC
GLM	AUC	1–1	NC	NC	0.97–1	1–1	NC	NC
—	TSS	0.99–1	NC	NC	0.94–1	1–1	NC	NC
maxent	AUC	0.98–0.99	0.88–0.93	1–1	0.99–0.99	0.99–0.99	0.99–1	0.99–1
—	TSS	0.69–0.94	0.68–0.8	0.77–0.98	0.78–0.92	0.77–0.9	0.86–0.9	0.88–0.97

**Table 1.** Range of the evaluation metrics for species distribution models calibrated with different climate datasets for *Xylella fastidiosa fastidiosa*, *X. fastidiosa multiplex* and *X. fastidiosa pauca*. For each algorithm and each dataset five models based on a subset of 80% randomly selected presence data were calibrated. We report the range of the metrics for models retained in the computation of the consensus models (TSS >0.6 and auc >0.85). bio6: minimum temperature of the coldest month; bio10: mean temperature of warmest quarter; bio11: mean temperature of the coldest quarter; bio15: precipitation seasonality. NC: not computed.

selected subset of eighty percent of the occurrence data (native and invaded ranges) and used the remaining twenty percent for model evaluation using AUC and TSS metrics. Five replications were done for each full model. Evaluations were done using a set of 100 randomly generated pseudo-absences.

We removed autocorrelated data to improve model predictive performances<sup>53,54</sup>. To do so, we used the first two axes of a Principal Component Analysis performed on the bioclimatic variables<sup>55</sup> recorded at each occurrence points. The first axis was divided into 100 bins. The bins of the second axis were fixed to have the same amplitude as for axis 1. The occurrence points were projected onto the resulting grid. When a grid cell contained more than one point, a random selection was used to retain only one point for further model calibration. This procedure was repeated for each *Xf* subspecies and for each climate dataset.

Model outputs were transformed into binary projections using the threshold that optimized the TSS statistics on the testing data<sup>56</sup>. The resulting projections were averaged to compute the committee (consensus) averaging that shows the likelihood of the presence of a species given the available data. The consensus model ranges from 0 (all the models predict absence) to 100 (all the models predict presence)<sup>44,56</sup>. We removed the individual models that did not reach the minimum quality threshold of TSS >0.6 and AUC >0.85<sup>56</sup> before computing the consensus (Table 1).

Our set of four algorithms comprised approaches belonging to the three main functional groups of species distribution models<sup>57</sup>. First, we selected the envelope model Bioclim<sup>58,59</sup> that relies only on presence data and as such makes no assumption about the absence of the organism under study<sup>60</sup>. Second we employed the maximum entropy algorithm Maxent that discriminates presences with background data<sup>61</sup>. Third, we used modeling techniques that rely on presence and absence or pseudo-absences namely the generalized linear model (GLM)<sup>62</sup> and the artificial neural network<sup>56,63</sup>. Maxent was fitted using 10,000 background points while GLM and artificial neural network were fitted using 200 pseudo-absences (Fig. S1).

Caution is usually warranted when interpreting models projected into new areas with climate conditions different from the calibration area<sup>64–66</sup>. Thus, we assessed the similarity of climate conditions between the calibration dataset and the projected area (*i.e.*, Europe) by computing the MESS index<sup>41</sup>.

The following R<sup>67</sup> packages were used to perform analyses and generate graphical outputs: biomod2<sup>68</sup>, cowplot<sup>69</sup>, dismo<sup>70</sup>, ecospat<sup>71</sup>, ggplot2<sup>72</sup>, raster<sup>73</sup> and rmaxent<sup>74</sup>.

## Results

**Algorithm and climate datasets selection.** The best combinations of algorithms  $\times$  climate datasets for *Xf multiplex* and *Xf pauca* were evaluated with a dataset comprising all European occurrences. As there were too few occurrences in Europe for *Xf fastidiosa*, a random subset of 20% of the native range occurrences was used. Models were selected on the basis of an arbitrary threshold of the TSS fixed to 0.6<sup>56</sup>. Only the following combination did not get sufficient statistical support to be retained in future analysis of the potential distribution of *Xf pauca*: bioclim  $\times$  dataset 2, 3, 6 and 7; GLM  $\times$  dataset 2, 3, 6 and 7 and Ann  $\times$  dataset 6 (Table 1). The remaining combinations were used to calibrate the full models based on the complete set of occurrences.

The accuracy of the resulting full models was examined and we only retained the full models associated to TSS > 0.6 and AUC > 0.85 for the computation of the consensus model. This quality check resulted in discarding seven, four and four models for *Xf fastidiosa*, *Xf multiplex* and *Xf pauca* respectively. Table 1 shows the range of the evaluation metrics for the models used to compute the consensus models (Fig. 2).

**Potential distribution of *Xylella fastidiosa*.** Figure 2 shows the committee averaging for the three subspecies of *Xf*. Each map shows the proportion of models predicting the presence of *Xf* in Western Europe. The potential distribution of *Xf* subsp. *fastidiosa* includes large regions of Spain, France, Italy, Croatia, Greece and Turkey as well as the coastal regions of North Africa (Fig. 2A). The proportion of models indicating favorable climate conditions in these areas was close or equal to 100%. The agreement between algorithms  $\times$  climate datasets was somewhat lower in the northern part of France, the British Islands, Belgium and Netherlands where a lower proportion of models predicted the presence of *Xf* subsp. *fastidiosa*. In the case of *Xf* subsp. *fastidiosa* (Fig. 2A), the MESS index indicated that climate conditions encountered within Western Europe do not differ from the climate conditions that characterize the native area (Supplementary Fig. S2).

The predicted potential distribution of *Xf* subsp. *multiplex* is depicted in Fig. 2B. European climate appears favorable in a very large area covering Spain, France, the British Isles, Italy, the Adriatic coast, Greece, Turkey and some coastal areas of the Black Sea. North Africa and Mediterranean coast of near East countries are also climatically suitable. The MESS index indicated a good match between the range of climate conditions that prevail in Europe and in the American range of *Xf* subsp. *multiplex* (Supplementary Fig. S3).

The predicted extent of climatically suitable conditions for *Xf* subsp. *pauca* is limited to the Mediterranean coastal regions with the exception of south Portugal and Spain Atlantic coasts. The MESS index indicated a mismatch between the minimum temperatures (bio6 and bio11) in northern Europe and the native area (Supplementary Fig. S4).

## Discussion

**Geographical distribution and possible impacts in Europe.** In a rapidly changing world, the design of pest control strategies (*e.g.*, early detection surveys and planning of phytosanitary measures) should ideally rely on accurate estimates of the potential distribution and/or impact of pest species as well as their responses to climate change<sup>75</sup>. In the present study, bioclimatic models predicted that a large part of the Mediterranean lowlands and Atlantic coastal areas of Europe are characterized by climatically suitable conditions for *Xf* subsp. *fastidiosa*, *multiplex* and *pauca* (Fig. 2). To a lower extent, favorable climate suitability is also observed in northern and eastern regions of Europe (North-eastern France, Belgium, the Netherlands, Germany, etc.).

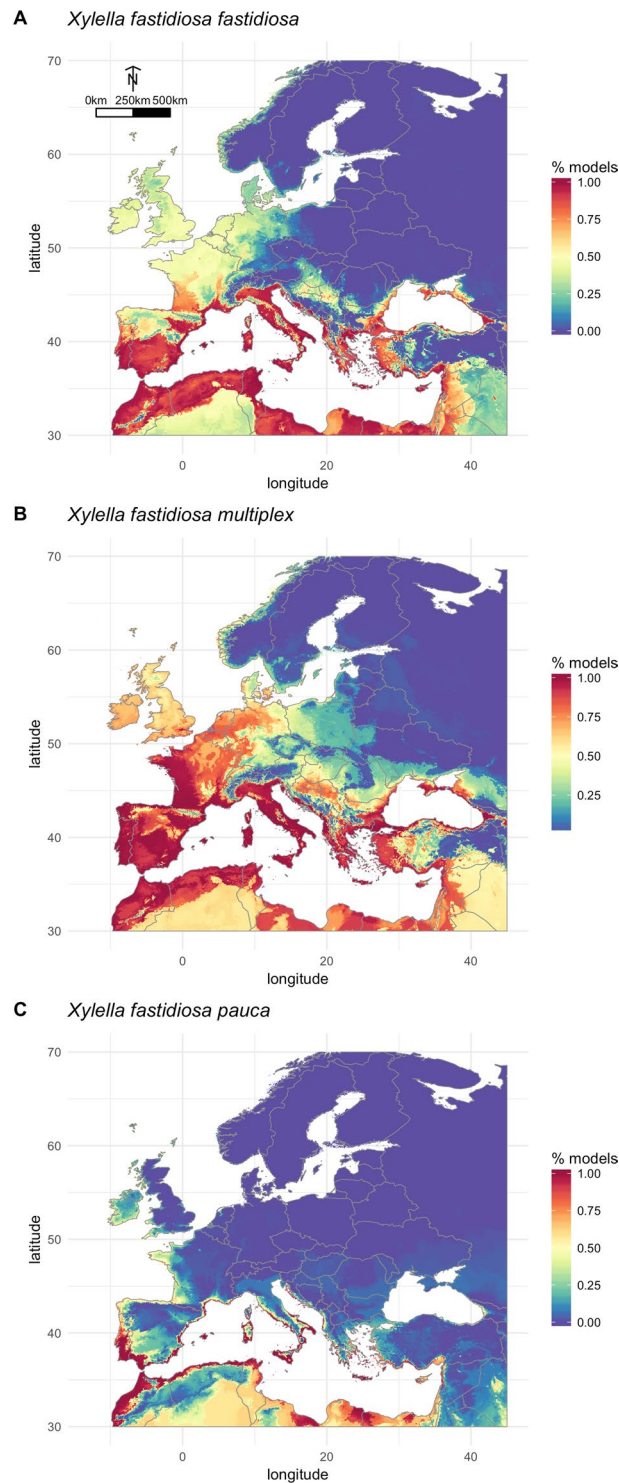
Our models displayed good evaluation measures and predicted high climatic suitability in all European areas where symptomatic plants are currently infected by the subspecies *fastidiosa*, *multiplex* or *pauca* (*e.g.*, Balearic Islands, lowlands of Corsica island, south-eastern France and the Apulia region). This suggests that climate suitability maps provided in the present study are reliable for the design of further sampling strategies, including ‘sentinel insects’ survey<sup>76,77</sup>. They may also be helpful to anticipate the spread of the different subspecies and provide guidance on which areas should be targeted for an analysis of local communities of potential vectors and host plants and to design further management strategies and research projects.

The results show that the different subspecies of *Xf* studied here might significantly expand in the near future, irrespective of climate change. For example, the subsp. *multiplex* known from Corsica and southern France have a large potential for expansion in Europe (Fig. 2B), which is not surprising since this subspecies has a wide distribution, ranging from Florida to Canada<sup>78</sup>. Actually, its expansion probably depends more on plant exchanges, vector spatio-temporal patterns and disease management than on climate suitability *per se*. *Xf multiplex* is associated to economically important plants such as almonds and olives<sup>26</sup> but may also colonize multiple ornamental plants or forest species<sup>27</sup>. Its potential distribution in Europe extends far beyond areas where the subspecies has been reported which suggests that new outbreaks may occur that could result in important economic losses.

The subspecies *fastidiosa*, which has been currently reported from a limited number of localities in Europe, could encounter favorable climate conditions in various areas (Fig. 2A). Notably, the models predict climate suitability in strategic wine-growing areas in different countries. The herein estimates of the potential distribution of the subsp. *fastidiosa* are consistent with the risk maps provided by Hoddle *et al.*<sup>39</sup> and Purcell (available in Anas *et al.*<sup>36</sup>).

The case of subsp. *pauca* is somewhat different (Fig. 2C). Most of the European occurrences come from southern Italy and the Balearic Islands and the potential distribution of this subspecies appears limited. Nevertheless,





**Figure 2.** Potential distribution of three subspecies of *Xylella fastidiosa*: (A) *Xf* subspecies *fastidiosa* (B) *Xf* subspecies *multiplex* (C) *Xf* subspecies *pauca*. Maps depict the ensemble forecast derived from committee averaging based on lowest presence thresholding (see methods section for details). The index varies from 100 when all models predict presence to 0 if all the models predict absence of the subspecies.

southern Spain and France, Portugal, Corsica, Sardinia, Sicilia and North Africa that are areas where growing olive trees is multisecular offer suitable conditions, which potentially implies huge socio-economic impacts.

One factor that proved to be critical for some insect-borne plant diseases is the distribution/availability of vectors and hosts. Here, none of these factors is limiting in most of Europe since *Xf* is capable of colonizing a vast array of plants present in Europe and *Philaenus spumarius*, the putatively most efficient European vector so far<sup>79,80</sup>, occurs across most of the continent<sup>77</sup>.

The potential distribution of the three subspecies of *Xf* studied here appeared to be limited by minimum winter temperatures with *Xf* subsp. *pauca* being much more sensible than the others. Because “cold curing” appears to be the main regulating mechanism, it is very likely that climate change would alter the distribution of suitable areas for *Xf* in Europe as the minimum winter temperatures might increase<sup>34,35,81</sup>. Furthermore, the potential dynamics of *Xf* in areas experiencing extremely high temperatures in summer (e.g., southern and central Spain) remain largely uncertain as the impact of extreme heat on *Xf* is poorly known<sup>82</sup>. Although warm spring and summer temperatures enhance multiplication of *Xf* in plants, it has been showed that *Xf* populations decrease in grapevines exposed to temperatures above 37 °C<sup>35</sup>. As southern and central Spain frequently experience temperatures above 40 °C in summer, additional data would be helpful to better understand the potential distribution and impact of *Xf* in these regions. Another point requiring clarification is the effect of rain and moisture. As *Xf* bacteria live in the xylem of plants they are subject to stress whenever their host is itself under water-stress. Precipitation or moisture may also have indirect impacts on *Xf* through insect vectors whose activity or behavior could be altered by water stress<sup>83</sup>. Although the three subspecies of *Xf* seem to have different tolerances to cold, it is, however, unclear whether realized niche divergence among subspecies reflects inherent differences in thermal tolerances or rather host-pathogen interactions as it was observed for *Ralstonia solanacearum*<sup>84</sup>. Additional investigations would allow a better understanding of the effect of temperatures on the different strains of *Xf*. It is noteworthy that potential distributions show large areas of potential co-occurrence. This may have important implications as it may increase the risk of intersubspecific homologous recombination<sup>11</sup>.

**Limits and opportunities for risk assessment.** Maps of habitat suitability should be cautiously interpreted as they are derived from correlative tools that depict the *realized* niche of species *i.e.*, a subset of the *fundamental* environmental tolerances constrained by biotic interactions, landscape structure and dispersal limits<sup>85</sup>. In addition, time-periods associated to occurrences and climate descriptors dataset do not perfectly overlap. The models were fitted with climate descriptors that represent average climate conditions for the 1970–2000 period, while some presence records were collected after 2000 in a period characterized by milder winter temperatures. Moreover, we deliberately fitted the models using a few climate descriptors to avoid model over-parameterization and/or extrapolation and enhance model transferability. Consequently, we cannot exclude that bioclimatic models presented here did not fully capture the entire range of environmental tolerances and did not fully depict the complexity of the climatic niche of *Xf* as well as potential interactions between climate descriptors. Better models and hence, better risk assessment could be obtained by collecting additional occurrence data as well as reliable absence data. The possible adaptation of *Xf* to new environmental constraints in its invaded range (e.g., by genetic recombination) is another important source of uncertainty. Finally, it is worth noting that bioclimatic models predict climatic suitability of a geographic region for *Xf* rather than a proper risk of *Xf*-induced disease incidence. To predict the proper severity of *Xf*-induced diseases in a given locality, statistical models should account for many additional factors playing a role in *Xf* epidemiology, including e.g., microclimate conditions, inter-annual climate variability, host-plant sensitivity, host-pathogen interactions, landscape structure and the spatio-temporal structure of the community of potential vectors. Although recent entomological studies identified the meadow spittlebug *P. spumarius* as the main vector of *Xf* in Italy<sup>79,80</sup>, a better knowledge of all European vectors capable of transmitting *Xf* to plants as well as their ecological characteristics (geographic range, efficiency in *Xf* transmission, demography, overwintering stage, intra-specific diversity, etc.) is needed<sup>86</sup>. In this context, habitat suitability maps could allow to design cost-efficient vector surveys, with priority given to geographic regions predicted as highly climatically suitable for *Xf*. The study by Cruaud *et al.*<sup>77</sup> provides a good insight into how species distribution modeling and DNA sequencing approaches may be combined for an accurate monitoring of the range of *Xf* and its vectors in Europe. We believe that SDMs are valuable tools to help in designing research experiments, control strategies as well as political decisions at the European scale.

## Conclusions/highlights

Species distribution models all indicate that the currently reported geographical range of *Xf* in Europe is small compared to the large extent of climatically suitable areas. This is true for all studied subspecies of *Xf* although the subspecies *pauca* appears to have a smaller potential range. *Xf* has a certain potential to adapt to climate and biotic conditions (hosts, vectors) encountered in Europe. The magnitude of this adaptive potential remains largely unknown but could nevertheless lead to a substantial spread of this plant pathogen across Europe. A further important research effort is thus needed to decipher the potential host plants – insect vectors – bacterium interactions in the (sub)natural ecosystems as well as agro-ecosystems at risk<sup>55</sup>. Only in this way could we develop an appropriate and efficient strategy to control *Xf* in the future.

## Data Availability

The occurrence data in France were made available to the authors by the French Ministère de l’Agriculture et de l’Alimentation subject to confidentiality requirements. The occurrence data in North and South America are available upon request to the authors.

## References

1. EFSA. Scientific opinion on the risks to plant health posed by *Xylella fastidiosa* in the EU territory, with the identification and evaluation of risk reduction options. *EFSA J* **13**, 3989 (2015).
2. Janse, J. & Obradovic, A. *Xylella fastidiosa*: its biology, diagnosis, control and risks. *J Plant Pathol* **92**, S35–S48 (2010).
3. Tumber, K., Alston, J. & Fuller, K. Pierce’s disease costs California \$104 million per year. *Calif Agric* **68**, 20–29 (2014).
4. Davis, M. J., Purcell, A. H. & Thomson, S. V. Pierce’s disease of grapevines: isolation of the causal bacterium. *Science* **199**, 75–77 (1978).
5. Martelli, G. P., Boscia, D., Porcelli, F. & Saponari, M. The olive quick decline syndrome in south-east Italy: a threatening phytosanitary emergency. *J Phytopathol* **144**, 235–243 (2016).

6. Hearon, S. S., Sherald, J. L. & Kostka, S. J. Association of xylem-limited bacteria with elm, sycamore, and oak leaf scorch. *Can J Bot* **58**, 1986–1993 (1980).
7. Wells, J. M., Raju, B. & Nyland, G. Isolation, culture and pathogenicity of the bacterium causing phony disease of peach. *Phytopathology* **73**, 859–862 (1983).
8. Chang, C. J., Garnier, M., Zreik, L., Rossetti, V. & Bové, J. M. Culture and serological detection of the xylem-limited bacterium causing citrus variegated chlorosis and its identification as a strain of *Xylella fastidiosa*. *Curr Microbiol* **27**, 137–142 (1993).
9. Davis, M., Thomson, S. & Purcell, A. Etiological role of a xylem-limited bacterium causing Pierce's disease in almond leaf scorch. *Phytopathology* **70**, 472–475 (1980).
10. Almeida, R. P. & Nunney, L. How do plant diseases caused by *Xylella fastidiosa* emerge? *Plant Dis* **99**, 1457–1467 (2015).
11. Scally, M., Schuenzel, E. L., Stouthamer, R. & Nunney, L. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. *Appl Environ Microbiol* **71**, 8491–8499 (2005).
12. Yuan, X. *et al.* Multilocus sequence typing of *Xylella fastidiosa* causing Pierce's disease and oleander leaf scorch in the United States. *Phytopathology* **100**, 601–611 (2010).
13. Nunney, L., Ortiz, B., Russell, S. A., Sánchez, R. R. & Stouthamer, R. The complex biogeography of the plant pathogen *Xylella fastidiosa*: genetic evidence of introductions and subspecific introgression in Central America. *PLoS One* **9**, e112463 (2014).
14. Nunney, L., Elfekih, S. & Stouthamer, R. The importance of multilocus sequence typing: cautionary tales from the bacterium *Xylella fastidiosa*. *Phytopathology* **102**, 456–460 (2012).
15. Schuenzel, E. L., Scally, M., Stouthamer, R. & Nunney, L. A multigene phylogenetic study of clonal diversity and divergence in North American strains of the plant pathogen *Xylella fastidiosa*. *Appl Environ Microbiol* **71**, 3832–3839 (2005).
16. Marcelletti, S. & Scortichini, M. Genome-wide comparison and taxonomic relatedness of multiple *Xylella fastidiosa* strains reveal the occurrence of three subspecies and a new *Xylella* species. *Arch Microbiol* **198**, 803–812 (2016).
17. Bull, C. *et al.* List of new names of plant pathogenic bacteria (2008–2010). *J Plant Pathol* **94**, 21–27 (2012).
18. Wells, J. M. *et al.* *Xylella fastidiosa* gen. nov., sp. nov.: gram-negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas* spp. *Int J Syst Evol Microbiol* **37**, 136–143 (1987).
19. Nunney, L. *et al.* Population genomic analysis of a bacterial plant pathogen: novel insight into the origin of Pierce's disease of grapevine in the US. *PLoS One* **5**, e15488 (2010).
20. Nunes, L. R. *et al.* Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements. *Genome Res* **13**, 570–578 (2003).
21. Coletta-Filho, H. D., Francisco, C. S., Lopes, J. R., Muller, C. & Almeida, R. P. Homologous recombination and *Xylella fastidiosa* host–pathogen associations in South America. *Phytopathology* **107**, 305–312 (2017).
22. De Lima, J. *et al.* Coffee leaf scorch bacterium: axenic culture, pathogenicity, and comparison with *Xylella fastidiosa* of citrus. *Plant Dis* **82**, 94–97 (1998).
23. Nunney, L., Yuan, X., Bromley, R. E. & Stouthamer, R. Detecting genetic introgression: high levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. *Appl Environ Microbiol* **78**, 4702–4714 (2012).
24. Nunney, L., Schuenzel, E. L., Scally, M., Bromley, R. E. & Stouthamer, R. Large-scale intersubspecific recombination in the plant-pathogenic bacterium *Xylella fastidiosa* is associated with the host shift to mulberry. *Appl Environ Microbiol* **80**, 3025–3033 (2014).
25. Randall, J. J. *et al.* Genetic analysis of a novel *Xylella fastidiosa* subspecies found in the southwestern United States. *Appl Environ Microbiol* **75**, 5631–5638 (2009).
26. Nunney, L. *et al.* Recent evolutionary radiation and host plant specialization in the *Xylella fastidiosa* subspecies native to the United States. *Appl Environ Microbiol* **79**, 2189–2200 (2013).
27. Denancé, N. *et al.* Several subspecies and sequence types are associated with the emergence of *Xylella fastidiosa* in natural settings in France. *Plant Pathol* **66**, 1054–1064 (2017).
28. EPPO. EPPO Standards. PM 7/24 (2) *Xylella fastidiosa*. EPPO bull **46**, 463–500 (2016).
29. Olmo, D. *et al.* First detection of *Xylella fastidiosa* on cherry (*Prunus avium*) and *Polygala myrtifolia* plants, in Mallorca Island, Spain. *Plant Dis* **101**, 1820 (2017).
30. Amanifar, N., Taghavi, M., Izadpanah, K. & Babaei, G. Isolation and pathogenicity of *Xylella fastidiosa* from grapevine and almond in Iran. *Phytopathol Mediterr* **53**, 318 (2014).
31. Çağlar, B. *et al.* First report of almond leaf scorch in Turkey. *J Plant Pathol* **87**, 246 (2005).
32. Su, C. C. *et al.* Pierce's disease of grapevines in Taiwan: Isolation, cultivation and pathogenicity of *Xylella fastidiosa*. *J Phytopathol* **161**, 389–396 (2013).
33. Wallingford, A. K., Myers, A. L. & Wolf, T. K. Expansion of the range of Pierce's disease in Virginia. *Online. Plant Health Progress*, <https://doi.org/10.1094/PHP-2007-1004-01-BR> (2007).
34. Purcell, A. Environmental therapy for Pierce's disease of grapevines. *Plant Dis* **64**, 388–390 (1980).
35. Feil, H. & Purcell, A. H. Temperature-dependent growth and survival of *Xylella fastidiosa* *in vitro* and in potted grapevines. *Plant Dis* **85**, 1230–1234 (2001).
36. Anas, O., Harrison, U. J., Brannen, P. M. & Sutton, T. B. Effect of warming winter temperatures on the severity of Pierce's disease in the Appalachian Mountains and Piedmont of the Southeastern United States. *Online. Plant Health Progress* **1**, 450–459, <https://doi.org/10.1094/PHP-2008-0718-01-RS> (2008).
37. Lieth, J., Meyer, M., Yeo, K.-H. & Kirkpatrick, B. Modeling cold curing of Pierce's disease in *Vitis vinifera* 'Pinot Noir' and 'Cabernet sauvignon' grapevines in California. *Phytopathology* **101**, 1492–1500 (2011).
38. Bosso, L., Febbraro, M., Cristinzio, G., Zoina, A. & Russo, D. Shedding light on the effects of climate change on the potential distribution of *Xylella fastidiosa*. *Biol Invasions* **18**, 1759–1768 (2016).
39. Hoddle, M. S. The potential adventive geographic range of glassy-winged sharpshooter, *Homalodisca coagulata* and the grape pathogen *Xylella fastidiosa*: implications for California and other grape growing regions of the world. *Crop Prot* **23**, 691–699 (2004).
40. VanDerWal, J., Shoo, L. P., Graham, C. & Williams, S. E. Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecol Model* **220**, 589–594 (2009).
41. Elith, J., Kearney, M. & Phillips, S. The art of modelling range-shifting species. *Methods Ecol Evol* **1**, 330–342 (2010).
42. Fick, S. E. & Hijmans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* **37**, 4302–4315 (2017).
43. Franklin, J. *Mapping species distributions: spatial inference and prediction*. (Cambridge University Press 2010).
44. Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R. K. & Thuiller, W. Evaluation of consensus methods in predictive species distribution modelling. *Divers Distrib* **15**, 59–69 (2009).
45. Zhu, G.-P. & Peterson, A. T. Do consensus models outperform individual models? Transferability evaluations of diverse modeling approaches for an invasive moth. *Biol Invasions*, 1–14 (2017).
46. Qiao, H., Soberón, J. & Peterson, A. T. No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. *Methods Ecol Evol* **6**, 1126–1136 (2015).
47. Araújo, M. B. & New, M. Ensemble forecasting of species distributions. *Trends Ecol Evol* **22**, 42–47 (2007).
48. Meller, L. *et al.* Ensemble distribution models in conservation prioritization: from consensus predictions to consensus reserve networks. *Divers Distrib* **20**, 309–321 (2014).
49. Jiménez-Valverde, A. *et al.* Use of niche models in invasive species risk assessments. *Biol Invasions* **13**, 2785–2797 (2011).



50. Martinetti, D. & Soubeyrand, S. Identifying lookouts for epidemics-surveillance: Application to the emergence of *Xylella fastidiosa* in France. *Phytopathology*, **109**, 265–276 (2019).
51. Liu, C., Berry, P. M., Dawson, T. P. & Pearson, R. G. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28**, 385–393 (2005).
52. Allouche, O., Tsoar, A. & Kadmon, R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol* **43**, 1223–1232 (2006).
53. Kramer-Schadt, S. *et al.* The importance of correcting for sampling bias in MaxEnt species distribution models. *Diversity and Distributions* **19**, 1366–1379 (2013).
54. Varela, S., Anderson, R. P., García-Valdés, R. & Fernández-González, F. Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography* **37**, 1084–1091 (2014).
55. Rasplus, J.-Y. *et al.* In AFPP – 4e conférence sur l'entretien des jardins végétalisés et infrastructures. (Available from, [http://arbestense.it/images/Annales\\_JEVI\\_2016.compressed.pdf](http://arbestense.it/images/Annales_JEVI_2016.compressed.pdf)).
56. Guisan, A., Thuiller, W. & Zimmermann, N. E. *Habitat suitability and distribution models*. (Cambridge University Press, 2017).
57. Peterson, A. T. *Ecological niches and geographic distributions*. (Princeton Univ. Press, 2011).
58. Busby, J. R. BIOCLIM: a bioclimate analysis and prediction system. *Plant Prot Q* **6**, 8–9 (1991).
59. Booth, T. H., Nix, H. A., Busby, J. R. & Hutchinson, M. F. BIOCLIM: the first species distribution modelling package, its early applications and relevance to most current MAXENT studies. *Biodivers Conserv* **20**, 1–9 (2014).
60. Dormann, C. F. Promising the future? Global change projections of species distributions. *Basic Appl Ecol* **8**, 387–397 (2007).
61. Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. *Ecol modelling* **190**, 231–259 (2006).
62. Guisan, A., Edwards, T. C. & Hastie, T. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecol Model* **157**, 89–100 (2002).
63. Ripley, B. D. *Pattern recognition and neural network*. (Cambridge University Press 1996).
64. Owens, H. L. *et al.* Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecol Model* **263**, 10–18 (2013).
65. Mesgaran, M. B., Cousens, R. D., Webber, B. L. & Franklin, J. Here be dragons: a tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. *Divers Distrib* **20**, 1147–1159 (2014).
66. Bell, D. M. & Schlaepfer, D. R. On the dangers of model complexity without ecological justification in species distribution modeling. *Ecol Model* **330**, 50–59 (2016).
67. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> (2017).
68. Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M. B. BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography* **32**, 369–373 (2009).
69. Wilke, C. O. cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. R package version 0.7.0, <https://CRAN.R-project.org/package=cowplot> (2016).
70. Hijmans, R. J., Phillips, S., Leathwick, J. & Elith, J. dismo: Species Distribution Modeling. R package version 1.1–4, <https://CRAN.R-project.org/package=dismo> (2017).
71. ecospat: Spatial Ecology Miscellaneous Methods. R package version 2.1.1. <https://CRAN.R-project.org/package=ecospat> (2016).
72. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag 2016).
73. Hijmans, R. J. & van Etten, J. raster: Geographic data analysis and modeling. R package version 2.6–7, <https://CRAN.R-project.org/package=raster> (2017).
74. Baumgartner, J. & Wilson, P. rmaxent: Tools for working with Maxent in R. Version 0.7.11.9, <https://github.com/johnbaums/rmaxent>.
75. Keller, R. P., Lodge, D. M. & Finnoff, D. C. Risk assessment for invasive species produces net bioeconomic benefits. *Proc Natl Acad Sci* **104**, 203–207 (2007).
76. Yaseen, T. *et al.* On-site detection of *Xylella fastidiosa* in host plants and in “spy insects” using the real-time loop-mediated isothermal amplification method. *Phytopathol Mediterr* **54**, 488–496 (2015).
77. Cruaud, A. *et al.* Using insects to detect, monitor and predict the distribution of *Xylella fastidiosa*: a case study in Corsica. *Scientific Reports* **8**, 15628 (2018).
78. Goodwin, P. & Zhang, S. Distribution of *Xylella fastidiosa* in southern Ontario as determined by the polymerase chain reaction. *Can J Plant Pathol* **19**, 13–18 (1997).
79. Cornara, D. *et al.* Transmission of *Xylella fastidiosa* by naturally infected *Philaenus spumarius* (Hemiptera, Aphrophoridae) to different host plants. *J Appl Entomol* **141**, 80–87 (2016).
80. Saponari, M. *et al.* Infectivity and transmission of *Xylella fastidiosa* by *Philaenus spumarius* (Hemiptera: Aphrophoridae) in Apulia, Italy. *J Econ Entomol* **107**, 1316–1319 (2014).
81. Purcell, A. Cold therapy of Pierce's disease of grapevines. *Plant Dis Repr* **61**, 514–518 (1977).
82. Daugherty, M. P., Zeilinger, A. R. & Almeida, R. P. P. Conflicting effects of climate and vector behavior on the spread of a plant pathogen. *Phytobiomes* **1**, 46–53 (2017).
83. Krugner, R. & Backus, E. A. Plant water stress effects on stylet probing behaviors of *Homalodisca vitripennis* (Hemiptera: Cicadellidae) associated with acquisition and inoculation of the bacterium *Xylella fastidiosa*. *J Econ Entomol* **107**, 66–74 (2014).
84. Milling, A., Meng, F., Denny, T. P. & Allen, C. Interactions with hosts at cool temperatures, not cold tolerance, explain the unique epidemiology of *Ralstonia solanacearum* race 3 biovar 2. *Phytopathology* **99**, 1127–1134 (2009).
85. Soberón, J. & Peterson, A. T. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodivers Inform* **2**, 1–10 (2005).
86. Chauvel, G., Cruaud, A., Legendre, B., Germain, J.-F. & Rasplus, J.-Y. Rapport de mission d'expertise sur *Xylella fastidiosa* en Corse. French Ministry of agriculture and food, Available at [http://agriculture.gouv.fr/sites/minagri/files/20150908\\_rapport\\_mission\\_corse\\_xylella\\_31082015b.pdf](http://agriculture.gouv.fr/sites/minagri/files/20150908_rapport_mission_corse_xylella_31082015b.pdf) (2015).

## Acknowledgements

We are indebted to the editor and three anonymous referees for their insightful comments and suggestions to this manuscript. We thank Pauline De Jerphanion (ANSES) manager of the French national database of *Xylella fastidiosa* in France as well as the DGAL for feeding that database. We also thank Christian Lannou (INRA, SPE, France) for his interest and support during the course of the project. This work was funded by grants from the SPE department of the INRA (National Agronomic Institute). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author Contributions

Designed the study: M.G., J.P.R.; performed the statistical analyses: M.G., J.P.R.; discussed the results: M.G., J.P.R., A.C., J.C.S. and J.Y.R., wrote the manuscript: M.G., J.P.R. All authors commented on the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-45365-y>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019